DOCUMENT RESUME

ED 276 756 TM 860 707

AUTHOR Kilgallon, Mary K.; Mueller, Richard J.
TITLE A Longitudinal Study of the Predictive Validity of a Kindergarten Screening Battery.
PUB DATE [86]
NOTE 33p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Kindergarten Children; *Learning Disabilities; Longitudinal Studies; *Predictive Validity; Predictor Variables; *Preschool Tests; Primary Education; Reading Comprehension; *Reading Tests; *Screening Tests; Sex Differences; *Test Validity
IDENTIFIERS *Test Batteries

ABSTRACT

Test validity was studied in nine subtests of a kindergarten screening battery used to predict reading comprehension for children up to five years after entering kindergarten. The independent variables were kindergarteners' scores on the: (1) Otis-Lennon Mental Ability Test; (2) Bender Visual Motor Gestalt Test; (3) Detroit Tests of Learning Aptitude (DTLA) Visual Attention Span for Objects; (4) DTLA Auditory Attention Span for Unrelated Words; (5) Elenbogen Test of Auditory Discrimination; (6) Word Recognition (deHirsch et al.); (7) Word Reproduction (deHirsch); (8) Wide Range Achievement Test (WRAT) reading achievement; and (9) WRAT arithmetic achievement. Gender and intelligence were predictor variables; test battery scores were compared with the dependent variable, the Stanford Achievement Test (SAT) reading comprehension subtest for grades one through five. Subjects were 245 suburban students. The Reading Comprehension subtest of the SAT was administered every spring through the fifth grade. Multiple regression analysis was performed for each grade and sex, generating prediction equations for SAT scores. Gender did not generally correlate significantly with other variables, including WRAT reading and mathematics. Results indicated significant relationships between the independent variables and reading comprehension. Stepwise regression determined the most important variables at each grade level. A bibliography and numerous tables are included. (GDC)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED276756

# A LONGITUDINAL STUDY OF THE PREDICTIVE VALIDITY OF A KINDERGARTEN SCREENING BATTERY

Mary K. Kilgallon
School Psychologist, District 67
Morton Grove, Illinois 60053

Richard J. Mueller
Dept of Learning, Development, & Special Education
Northern Illinois University
DeKalb, Illinois 60115

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. J. Mueller

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM 860 707

# A LONGITUDINAL STUDY OF THE PREDICTIVE VALIDITY OF A KINDERGARTEN SCREENING BATTERY

In recent years, a major area of concern in education has been the early identification of children who experience or are likely to experience learning disabilities later in school (Lindsay & Wedell, 1982). This interest is international in scope, as evidenced by a survey which showed that 74 percent of the Local Education Authorities in England and Wales either were using (47%), or planned to use (27%), screening procedures (Lewis, 1980).

The two procedures most frequently used are standardized readiness tests and teacher-completed rating scales. While both are helpful, neither has been completely satisfactory. Nurss (1979) lamented that "most readiness tests are group paper-and-pencil tests designed to be given by the classroom teacher within a reasonable amount of time. These limitations considerably reduce the variety of test formats and content that might be included" (p. 33). According to Nurss, correlations from .54 to .70 are typical, indicating that readiness tests do predict reading achievement; yet, "when considering individual children, the prediction may be quite imperfect" (p. 39).

Teacher-completed rating scales involve subjective judgement, and prediction appears to be even less. Lederman and Braine (1972) reported that teachers' ratings of children's readiness accounted for about 50 percent of the variance. Another authority, Keough, criticized teacher ratings of pupil readiness as putting too much emphasis on behavioral characteristics (1977).

> A major finding appears to be that there are at least two partially independent components of risk in kindergarten and first grade: The first is related to academic aptitude, the second to behavioral adaptability. . . . although children identified by teachers and by our measures as being at risk in kindergarten tended to be children with problems in behavioral adaptability, risk status in first grade was more likely to be associated with problems of academic aptitude (p. 270).

## PURPOSE OF THE STUDY

This study was undertaken to determine the validity of nine subtests of a kindergarten screening battery in predicting reading comprehension for children for up to five years after entering kindergarten. The independent variables for this investigation were (1) intelligence, (2) visual-motor integration, (3) short-term visual memory, (4) short-term auditory memory, (5) auditory discrimination, (6) reading achievement, (7) arithmetic achievement, (8) word recognition, and (9) word reproduction. Battery scores were compared with a measure of reading comprehension--the Stanford Achievement Test--for grades one through five. Intelligence is used as a predictor variable only. Gender is utilized as an assigned,

categorical variable as well as a predictor variable.

The subjects for this study were 245 pupils who attended kindergarten in two elementary schools located in a Chicago suburban community during the 1978 through 1982 school years. In the spring of each year through 1983, the children were administered the reading comprehension subtest of the SAT. Table 1 indicates the number of children in the kindergarten through fifth grade groups.

Table 1

Distribution of Kindergarten Screening Battery and Five Grades of Achievement Testing

| Year | Kindergarten | First Grade | Second Grade | Third Grade | Fourth Grade | Fifth Grade |
|---|---|---|---|---|---|---|
| 1978 | 53 | | | | | |
| 1979 | 60 | 50 | | | | |
| 1980 | 53 | 56 | 46 | | | |
| 1981 | 47 | 51 | 51 | 42 | | |
| 1982 | 44 | 45 | 49 | 49 | 39 | |
| 1983 | | 43 | 44 | 42 | 46 | 38 |
| | 257 | 245 | 190 | 133 | 85 | 38 |
| | | (95%) | (89%) | (80%) | (75%) | (72%) |

The study utilizes computerized, multiple regression statistical packages that provide opportunities for the analysis of data in several ways for the most advantageous results. The research design is multivariate, which is a statistical procedure of relatively recent origin in the prediction of early reading disabilities (Pope, Lehrer, & Stevens, 1980). Finally, there is a need for research on this subject (Lewis, 1980; Lindsay & Wedell, 1982).

## MULTIPLE REGRESSION

Regression models explain variance in dependent variables as a result of their relationships with independent variables. In these models, error is the difference between the predicted and the observed dependent variable scores. Regression analysis models are designed to minimize this error, as well as explain the variance itself.

Multiple regression yields three outcomes: (1) The multiple coefficient of correlation--R--which is the relationship between the dependent variable and the linear composite of independent variables; (2) the multiple coefficient of determination--R2--which is the proportion of dependent variable variance explained by its relationship to the linear composite of independent variables; and, (3) partial beta weights, which represent the relative importance of each independent variable as a predictor.

Multiple regression analyses from the Statistical Analysis System (Service, 1972) were used to generate prediction equations for the Reading Comprehension subtest of the Stanford Achievement Test, based upon the combinations of the screening battery. Regression analyses were performed for each grade and for the gender groups within each grade.

Pedhazur (1982) has noted that the variables under investigation in much of behavioral science research are often intercorrelated, and it is appropriate to determine if a combination of fewer variables will result in prediction as efficient or almost as efficient as the entire set. The intention is to ferret out the "best" predictors--the minimum number of variables possible that will account for as much, or almost as much of the variance accounted by the entire set (p. 151). It should be emphasized, however, that determining a truly "best" prediction is as much a conceptual process as it is statistical.

The stepwise procedure that was followed in this analysis provides for five methods (SAS). Each step is a test that determines the contribution of each predictor. Model selection routine utilized is MAXR (Barr, et al., 1976), which involves forward selection with pair switching. MAXR is the maximum R2 improvement technique developed by Goodnight (Pedhazur, 1982) and is considered to be almost as good as all possible regressions. The method is basically a multivariable model, and it will isolate the best one-variable model, the best two-variable model, and so on. However, it does not guarantee to find the model with the largest R2 for each size.

## REVIEW OF RELATED LITERATURE

Research relating to kindergarten screening contains numerous studies of test batteries developed to evaluate young children's skills and abilities (deHirsch, et al, 1966; Jansky & deHirsch, 1972; Book, 1974; Satz & Friel, 1978; Hillerich, 1978; Rubin, et al, 1978).

DeHirsch, Jansky, and Langford (1966) designed a study to identify kindergarten pupils who were expected to fail reading even

if they had average intelligence. The children were given a battery of 36 subtests and then tested again at the end of second grade. The result was a Predictive Index of 10 subtests (out of the original 36) that appear to most effectively identify high-risk children. However, because of the time-consuming nature of this first Predictive Index, Jansky and deHirsch (1972) developed a shorter screening battery of 6 subtests intended to measure large populations quickly and individually.

Book (1974) developed a predictive battery for the purposes of early identification and educational planning, consisting of the Metropolitan Readiness Test (1964), the Bender Gestalt Test (1938), and the Slosson Intelligence Test (1963). First, all kindergarten children were assessed with the Metropolitan Readiness Tests (1964). Then, if a child did poorly on this test, the rest of the battery was given. As a result, each child was placed into 6 categories for purposes of program planning.

Satz and Friel (1978) administered an abbreviated test battery at the beginning of kindergarten to determine if reading group membership can be predicted at the end of second grade. This abbreviated battery consisted of 8 tests which tapped 3 factors: sensorimotor-perceptual skills, verbal-conceptual skills, and verbal-cultural experiences. The results showed that the battery consistently identified over 90 percent of the children destined to become severely or superior readers in later years, although prediction was lower for children in the mid-range of achievement.

Hillerich's 1978 screening battery is called the Prediction with Diagnostic Qualities (PDQ), consisting of 9 subtests intended to

measure receptive and expressive language. Hillerich (1978) administered the PDQ to 916 beginning kindergarden students in a study having both experimental and control groups. Reading achievement tests administered at the end of first grade showed a significant difference in favor of the experimental group.

Rubin, et al. (1978) designed a study to investigate if children identified on readiness test as "high risk" do, in fact, achieve poorly by later measures. The Metropolitan Readiness Test (1964) was administered prior to kindergarten, again prior to first grade, and then compared to Stanford Achievement Test (1973) results at age 9. The results were significant correlations ranging from .50 to .71 between the Metropolitan scores and SAT achievement scores in reading, spelling, and math. Apparently, prediction is best for subsequent high achievers, but less accurate for the low achievers.

## INSTRUMENTATION

The deviation intelligence quotient obtained on the Otis-Lennon Mental Ability Test, Form J, Primary 1 Level was used as a measure of the learning ability of the children in this study. As with the previous editions of the Otis series, this test was constructed to a yield a dependable measure of "g", or general intelligence (Otis & Lennon, 1967). The items on the Otis-Lennon sample comprise the mental processes of classification, following directions, quantitative reasoning, and comprehension of verbal concepts.

Reviews of the Otis-Lennon by Mulholland (1972) and Smith (1970) agree that the 1967 edition adheres to the highest level of current test standards. According to Smith (1970), the reliability data will

give the user confidence in using the results of the Otis-Lennon. The Test Manual (Otis & Lennon, 1967) reports a Kuder-Richardson Number 21 correlation of .88 for the Primary 1 IQ test.

Visual Motor Integration

The Bender Visual Motor Gestalt Test (Bender, 1938) is one the most widely used psychological tests and the most popular of the visual motor tests (Sattler, 1982). Probably the most popular objective scoring system for the Bender Gestalt is the Koppitz Developmental Bender Scoring System (Koppitz, 1964). There are 30 developmental scoring items with each item receiving 1 or 0 points, depending on whether an error occurs. The points are summed to obtain a total score which is compared to norms for the child's age.

Test-retest reliabilities for the Developmental Scoring system have ranged from .50 to .90 with a median correlation of .77 (Koppitz, 1975). Concurrent validity for the Koppitz system was established by correlating it with various tests of visual-motor perception (Sattler, 1982).

Reading and Arithmetic Achievement

The Wide Range Achievement Test (WRAT) developed by Jastak and Jastak (1965 edition) was used to assess reading and arithmetic achievement. The WRAT is a brief, individually administered word recognition test which measures recognizing and naming letters and pronouncing words. Arithmetic skills in the Oral Arithmetic section include counting, reading number symbols, and solving oral problems. Sattler (1982) comments that "the WRAT. . . appears to be a useful screening device when a quick estimate of reading, arithmetic and spelling skills is needed" (p. 259).

The WRAT manual reports split-half reliabilities that range from . 94 to .98 for the subtests. As to validity, there have been many studies that compare the WRAT with other tests of achievement, intelligence, and ability. Sattler (1982) concludes "the results of these studies suggest that the WRAT has adequate concurrent validity" (p. 259).

### Short-term Visual and Auditory Memory

The Detroit Tests of Learning Aptitude (Baker & Leland, 1967) contains 19 subtests that measure various abilities--two of which were used in the present study. Visual Attention Span for Objects was used to assess short-term visual memory, and the Auditory Attention Span for Unrelated Words measures short-term auditory memory.

Test-retest reliability is reported in the manual and ranges from .96 for a small sample retested at an interval of five months to .68 for 792 students retested at an interval of two to three years (Sattler, 1982. Validity is cited terms of intercorrelations reported in the manual. These range between .20 and .40, which indicate relative independence of the subtests (Sattler, 1982).

### Auditory Discrimination

The Elenbogen Test of Auditory Discrimination is an altered form of the Wepman Test of Auditory Discrimination (1973), using nonsense syllables instead of real words as on the Wepman. However, Thompson (1972) found that high socio-economic status students do significantly better on the Wepman than on the Elenbogen--a limitation when testing children with enriched language skills.

### Word Recognition and Word Reproduction

The Word Recognition Test in the kindergarten screening battery is a combination of Word Recognition I and Word Recognition II as described by deHirsch, Jansky, and Langford (1966). Word Reproduction is also adopted from the deHirsch battery (1966).

## RESULTS

The dependent variable was the Reading Comprehension subtest of the Stanford Achievement Test. In grades one, two, and three, the Reading Comprehension test correlated significantly with all the other variables except gender. The Reading Comprehension test correlated with all variables except gender and the Bender Gestalt at grade four. The fifth grade Reading Comprehension correlations were also significant, with the exception of gender, IQ, the Elenbogen Test of Auditory Discrimination and Word Reproduction.

With the exception of I.Q. in second and the Detroit Visual Memory in fifth grades, gender did not correlate significantly with other variables. The Math subtest of the Wide Range Achievement Test (WRAT) correlated with all of the other variables except gender. The WRAT Reading subtest correlated with all variables in the first grade except gender, and in second through fourth grades was related to all variables except the Bender Gestalt and the Detroit Auditory Memory.

Intelligence, as measured by the Otis-Lennon Mental Ability Test, was significantly related to all of the other variables in grades one and two (except gender). However, IQ did not correlate with the Bender Gestalt or the Word Recognition test in either grades three or four.

The Detroit Visual Memory correlated with the other variables

(but not gender) at grades one and three. In grades two and four, the Detroit Visual Memory test correlated with all of the variables except gender and the Bender Gestalt.

With the exception of gender, Word Reproduction was correlated significantly with all the other variables in first and second grades. In grade three, Word Reproduction did not correlate significantly with either gender or the Detroit Auditory Memory test. At fourth grade, Word Reproduction also did not correlate with the Bender Gestalt, in addition to the exceptions noted above.

The Elenbogen Test of Auditory Discrimination correlated significantly with all the other variables except gender in grades one and two. In third grade, there was an insignificant relationship with the Detroit Auditory Memory test. Likewise, in fourth grade the Elenbogen did not correlate with Word Recognition. (Because the Elenbogen is scored for errors, the reported correlations have a negative sign).

Word Recognition correlated significantly with all the other variables except gender in first grade. In second grade, however, Word Recognition showed no relationships with either gender or the Detroit Auditory Memory test.

The Bender Gestalt and the Detroit Auditory Memory test revealed fewer significant correlations than the other variables. The Bender Gestalt is scored for errors, and this results in negative correlations.

By fifth grade, there were only 24 of a possible 55 significant correlations. Only WRAT Math was significantly correlated with all of the other variables except gender.

In summary, this analysis resulted in a large number of

significant relationships between the independent variables of the kindergarten screening battery and the dependent variable--reading comprehension.

## REGRESSION ANALYSES

A stepwise regression analysis was computed to determine the ranking of the predictor variables. An alpha of .05 was set, and only those variables whose beta weights were significant at the .05 level are reported.

There was a declining number of variables in the regression tables from grades one to five. Grade one had 6 variables, grade two had 5 variables, grade three had 3 variables, grade four had 2 variables and grade five had 1 variable.

At first grade, WRAT Mat entered the analysis first and explained 26 percent of the variance. Word Reproduction contributed 9 percent more to the variance. Word Recognition contributed 5 percent more; the Bender Gestalt added 3 percent; the Elenbogen Test of Auditory Discrimination added 2 percent; and the group I.Q. scores added 1 percent. Altogether, the 6 tests in the regression model accounted for 46 percent of the variance of the dependent variable--SAT Reading Comprehension ($p < .0001$).

Word Reproduction entered first into the regression model and accounted for 22 percent of the variance. I.Q. entered next and took 6 percent more. Adding the Elenbogen Test of Auditory Discrimination, WRAT Math, and the Detroit Test of Visual Memory contributed 9 percent more to the explanation of the variance. In all, the five tests in the regression model accounted for 37 percent of the variance of the Reading Comprehension subtest of the SAT ($p < .0001$).

The regression analysis for grade three contains 3 variables and accounted for 42 percent of the variance of the SAT Reading Comprehension subtest ($p < .0001$). Word Reproduction entered first into the regression model and accounted for 25 percent of the variance. I.Q. entered next and added 10 percent to the variance. The third variable--Word Recognition--added 7 percent to the variance.

At fourth grade, there were only two variables that fulfill the alpha=.05 condition, but these two tests accounted for 27 percent of the variance ($p < .0001$). I.Q. entered first and took 18 percent of the variance. The inclusion of Word Recognition in the model added another 7 percent to the variance.

The fifth grade regression analysis had only one variable---Detroit Visual Memory test--which accounted for 38 percent of the variance ($p < .0001$).

Since some of the subtests contributed significantly to the regression model, it would seem that there is a sufficiently clear pattern of subtests to warrant a shorter battery. A partial model, however, is not appropriate in this analysis because all 10 variables were used, and two of the subtests (Word Recognition and Word Reproduction) have a time component. Also, two words are taught to the children at the beginning of the testing situation. The children are then given the Elenbogen Test of Auditory Discrimination, Detroit Test of Visual Memory, and the Detroit Test of Auditory Memory. Following the administration of these tests, the Word Recognition and Word Reproduction tests are given (involving the two words previously taught). The child must retain these two words through the time span of these three intervening tests. Shortening

the battery will eliminate these intervening tasks, and thus the battery will not be administered under the conditions described in this study. Therefore, Tables 2, 3, 4, 5, and 6 identify the "best" 10-variable model for each grade selected by the MAXR, stepwise regression technique, and include the 13 significant beta weights in the full model (10 variables) regression results for each grade level.

Table 2

First Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)—Best 10-Variable Model R2=.465**

| Order of Entering Predictor Variable | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehension Test (Stanford Achievement Test) | | | | |
|---|---|---|---|---|---|---|
| | | R | R2 | R2 Change | B | F |
| 1. | WRAT Math | .505 | .255 | | 4.210 | 14.47** |
| 2. | Word Reproduction | .590 | .348 | .093 | .772 | 6.10* |
| 3. | Word Recognition | .630 | .397 | .049 | 2.093 | 17.53** |
| 4. | Bender Gestalt | .653 | .426 | .029 | -.339 | 8.27** |
| 5. | Elenbogen Aud. Disc. | .667 | .445 | .019 | -.264 | 7.30** |
| 6. | Intelligence Quotient | .675 | .455 | .010 | .060 | 3.31 |
| 7. | Detroit Visual Memory | .679 | .461 | .006 | .462 | 3.47 |
| 8. | Detroit Aud. Memory | .681 | .464 | .003 | -.334 | 1.54 |
| 9. | WRAT Reading | .682 | .465 | .001 | .535 | .61 |
| 10. | Sex | .682 | .465 | .000 | -.108 | .02 |

*p <.05

**p <.01

Table 3

Second Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)—Best 10-Variable Model R2=.387**

| Order of Entering Predictor Variable | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehension Test (Stanford Achievement Test) R | R2 | R2 Change | B | F |
|---|---|---|---|---|---|---|
| 1. | Word Reproduction | .469 | .220 | | .758 | 5.54* |
| 2. | Intelligence Quotient | .533 | .284 | .064 | .074 | 4.80* |
| 3. | Elenbogen Aud. Disc. | .569 | .324 | .040 | -.247 | 4.65* |
| 4. | WRAT Math | .590 | .349 | .025 | 1.876 | 2.51 |
| 5. | Detroit Vis. Memory | .604 | .365 | .016 | .412 | 2.28 |
| 6. | Word Recognition | .612 | .375 | .010 | .852 | 2.54 |
| 7. | WRAT Reading | .619 | .383 | .008 | 1.219 | 2.52 |
| 8. | Detroit Aud. Memory | .621 | .385 | .002 | .218 | .58 |
| 9. | Bender Gestalt | .622 | .387 | .002 | -.090 | .52 |
| 10. | Sex | .622 | .387 | .000 | .052 | .00 |

*p <.05

**p <.01

Table 4

Third Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)—Best 10-Variable Model R2=.449**

| Order of Entering Predictor Variable | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehension Test (Stanford Achievement Test) R | R2 | R2 Change | B | F |
|---|---|---|---|---|---|---|
| 1. | Word Reproduction | .495 | .245 | | 1.452 | 8.15** |
| 2. | Intelligence Quotient | .590 | .348 | .076 | .168 | 10.35** |
| 3. | Word Recognition | .651 | .424 | .012 | 2.862 | 12.95** |
| 4. | WRAT Reading | .660 | .436 | .008 | 1.452 | 1.90 |
| 5. | Detroit Aud. Memory | .666 | .444 | .003 | .502 | 1.28 |
| 6. | Bender Gestalt | .668 | .447 | .001 | -.139 | .54 |
| 7. | WRAT Math | .669 | .448 | .001 | .918 | .25 |
| 8. | Sex | .670 | .449 | .000 | .315 | .06 |
| 9. | Detroit Visual Memory | .670 | .449 | .000 | .061 | .02 |
| 10. | Elenbogen Aud. Disc. | .670 | .449 | .000 | .017 | .01 |

*p <.05

**p <.01

Table 5

Fourth Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)—Best 10-Variable Model R2=.342**

| Order of Entering Predictor Variable | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehension Test (Stanford Achievement Test) | | | | |
|---|---|---|---|---|---|---|
| | | R | R2 | R2 Change | B | F |
| 1. | Intelligence Quotient | .429 | .184 | | .296 | 6.96* |
| 2. | Word Recognition | .506 | .256 | .072 | 3.592 | 4.50* |
| 3. | Word Reproduction | .531 | .282 | .026 | 1.085 | .98 |
| 4. | Sex | .551 | .304 | .022 | 5.091 | 3.32 |
| 5. | Detroit Aud. Memory | .563 | .317 | .013 | .847 | .78 |
| 6. | Bender Gestalt | .573 | .328 | .011 | .536 | 1.62 |
| 7. | WRAT Math | .581 | .358 | .010 | 3.969 | .64 |
| 8. | WRAT Reading | .584 | .341 | .003 | 1.220 | .30 |
| 9. | Detroit Visual Memory | .585 | .342 | .001 | .243 | .05 |
| 10. | Elenbogen Aud. Disc. | .585 | .342 | .000 | -.064 | .02 |

*p <.05

**p <.01

Table 6

Fifth Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)--Best 10-Variable Model R2=.513*

| Order of Entering Predictor Variable | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehension Test (Stanford Achievement Test) | | | | |
|---|---|---|---|---|---|---|
| | | R | R2 | R2 Change | B | F |
| 1. | Detroit Visual Memory | .619 | .383 | | 3.386 | 2.33 |
| 2. | Word Recognition | .661 | .437 | .054 | 2.528 | 1.02 |
| 3. | WRAT Math | .677 | .459 | .022 | 5.548 | .59 |
| 4. | Detroit Aud. Memory | .688 | .474 | .015 | 1.035 | .41 |
| 5. | Elenbogen Aud. Disc. | .696 | .485 | .011 | .728 | 1.25 |
| 6. | Bender Gestalt | .702 | .493 | .003 | -.776 | 1.27 |
| 7. | WRAT Reading | .713 | .509 | .016 | 1.652 | .22 |
| 8. | Sex | .716 | .512 | .000 | -.723 | .16 |
| 9. | Intelligence Quotient | .716 | .512 | .000 | -.046 | .04 |
| 10. | Word Reproduction | .716 | .513 | .001 | -.174 | .01 |

*p <.05

**p <.01

If there are no restrictions on the regression models, the 10 variables of the kindergarten screening battery will explain 47 percent of the variance in grade one, 39 percent in grade two, 45 percent in grade three, 34 percent in grade four, and 51 percent in grade five.

Tables 7 through 11 indicate the abbreviated regression models for the five grades. The multiple R's were significant ($p < .05$). The tables indicate the 17 significant beta weights ($p < .05$) in the abbreviated models. Six of these beta weights were at first grade (Table 7), 5 at second grade (Table 8), 3 at grade three (Table 9), 2 at grade four (Table 10), and 1 at grade five (Table 11).

Table 7

First Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)

| Order of Entering Predictor Variable | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehension Test (Stanford Achievement Test) | | | | |
|---|---|---|---|---|---|---|
| | | R | R2 | R2 Change | B | F |
| 1. | WRAT Math | .505 | .255 | | 4.560 | 18.78** |
| 2. | Word Reproduction | .590 | .348 | .093 | .942 | 10.13** |
| 3. | Word Recognition | .630 | .397 | .049 | 2.119 | 18.04** |
| 4. | Bender Gestalt | .653 | .426 | .026 | -.346 | 8.62** |
| 5. | Elenbogen Aud. Disc. | .667 | .445 | .019 | -.269 | 7.93** |
| 6. | Intelligence Quotient | .675 | .455 | .010 | .068 | 4.48* |

*p <.05

**p <.01

Table 8

Second Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)

| Order of Entering Predictor Variable | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehsion Test (Stanford Achievement Test) | | | | |
|---|---|---|---|---|---|---|
| | | R | R2 | R2 Change | B | F |
| 1. | Word Reproduction | .469 | .220 | | .998 | 10.97** |
| 2. | Intelligence Quotient | .533 | .284 | .064 | .082 | 6.05* |
| 3. | Elenbogen Aud. Disc. | .569 | .324 | .040 | -.313 | 7.98** |
| 4. | Wide Range Math | .591 | .349 | .025 | 2.525 | 4.76* |
| 5. | Detroit Visual Memory | .604 | .365 | .016 | .538 | 4.64* |

*p <.05

**p <.01

Table 9

Third Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)

| Order of Entering Predictor Variables | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehension Test (Stanford Achievement Test) | | | | |
|---|---|---|---|---|---|---|
| | | R | R2 | R2 Change | B | F |
| 1. | Word Reproduction | .495 | .245 | | 1.860 | 18.52** |
| 2. | Intelligence Quotient | .590 | .348 | .103 | .206 | 19.59** |
| 3. | Word Recognition | .651 | .424 | .076 | 3.157 | 17.10** |

*p <.05

**p <.01

Table 10

Fourth Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)

| Order of Entering Predictor Variables | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehension Test (Stanford Achievement Test) | | | | |
|---|---|---|---|---|---|---|
| | | R | R2 | R2 Change | B | F |
| 1. | Intelligence Quotient | .429 | .184 | | .415 | 16.29** |
| 2. | Word Recognition | .506 | .256 | .072 | 4.449 | 7.94** |

*p <.05

**p <.01

Table 11

Fifth Grade Reading Achievement: Stepwise Multiple Regression Analysis (Maximum R2 Improvement)

| Order of Entering Predictor Variables | Predictor Variables: Kindergarten Screening Battery Tests | Criterion Variable: Reading Comprehension Test (Stanford Achievement Test) | | | | |
|---|---|---|---|---|---|---|
| | | R | R2 | R2 Change | B | F |
| 1. | Detroit Visual Memory | .619 | .383 | | 5.166 | 22.32** |

*p <.05

**p <.01

An hypothesis of this study was concerned with gender as a predictor of reading comprehension and whether there should be a separate regression table for boys and girls. Gender resulted in only one significant correlation in the analysis of the independent variables of the kindergarten screening battery. The correlations between gender and the dependent variable--reading comprehension--were not significant. Consequently, there appears to be no need for separate regression equations for boys and girls.

In summary, the analysis of the data indicate that there are significant relationships among the subtests of the kindergarten screening battery with the dependent variable--reading comprehension. The multiple regression equations based upon the kindergarten screening battery does predict reading comprehension for the whole group and for the subgroups. The subtests of the kindergarten screening battery make significant contributions to the prediction of reading comprehension. Gender, however, was not a significant predictor of reading comprehension in this study.

## CONCLUSION

The intent of this study was to determine if the present instruments are effective in the early identification of children who may have later difficulties, and thus provide a basis for appropriate diagnosis and intervention. The screening test does, in fact, have a high degree of predictive accuracy. In addition, the data suggest that critical developmental characteristics already exist in measurable amounts early in the school years. It appears that there is not a primary lag or deficit in any single modality or function, (e.g., auditory sequencing, memory, etc.). The evidence points, instead, to

a general developmental lag or advance in the behaviors measured by the battery subtests.

Another finding of this study is that there appears to be no single skill area or function deficiency that stands out above all others. In light of this, there is little justification for singling out faulty reading instruction as the primary cause of reading failure. The relationship between deficiencies identified at the kindergarten stage and subsequent reading skills appears to be a complex one. The onset of reading difficulties does not fit the pattern of a clear, causal link, as, for example, between a measurable hearing loss and delayed speech.

An interesting--though somewhat unexpected--outcome of this study is the lack of gender differences in the prediction equation. This supports the view by many that gender differences in reading achievement in regular classrooms--in the early years--is not as great as supposed (Lambert & Nicholl, 1976).

The investigator believes that the following uses for the kindergarten screening battery are appropriate:

1. The instrument may serve as a first-step, broad screening procedure, to be followed by more intensive study and assessment of children who manifest problems in their kindergarten performance.

2. The instrument can be used by the classroom teacher for systematic observation procedures, identification of specific areas of learning difficulties, and a basis for individual, prescriptive objectives.

3. Since the screening of kindergarten children provides fairly accurate predictions of their reading achievement through the fifth

grade, the battery can provide the basis for continuous analysis and revision of the instructional program. Supplementary and alternative instructional programs can be implemented as early as first grade.

There is little question that there is a growing recognition and acceptance of the belief that a full and appropriate education is a right for all children, including those with handicaps, developmental delays, or learning problems. Valid and accurate diagnosis in the early months of a child's educational experiences could be a major determinant of his/her achievement throughout all of the school years to come.

## REFERENCES

Baker, H. J., & Leland, B. (1967). *Detroit Tests of Learning Aptitude--Examination Manual*. Indianapolis: Bobbs-Merrill.

Barr, A. J., Goodnight, J. H., Sall, J. P., & Helwig, J. T. (1976). *SAS: A user's guide to SAS 76*. Raleigh, N. C.: SAS Institute.

Bender, L. (1938). A visual-motor gestalt test and its clinical use. *American Orthopsychiatric Association Research Monograph*.

Book, R. M. (1974). Predicting reading failure: A screening battery for kindergarten children. *Journal of Learning Disabilities, 7*(1), 52-56.

deHirsch, K., Jansky, J., & Langford, W. S. (1966). *Predicting reading failure*. New York: Harper & Row.

Hillerich, R. L. (1978). A diagnostic approach to early identification of languager skills. *The Reading Teacher, 31*(4), 357-364.

Jansky, J., & deHirsch, K. (1972). *Preventing reading failure.* New York: Harper & Row.

Keough, B. K. (1977). Early identification: Selective perception or perceptive selection? *Academic Therapy, 2*(3), 267-274.

Koppitz, E. M. (1964). *The Bender-Gestalt Test for Young Children: Research and application (vol. 2), 1963-1973.* New York: Grune & Stratton.

Lambert, N. M., & Nicholl, R. C. (1976). Dimensions of adaptive behavior of retarded and non-retarded public school children. *American Journal of Mental Deficiency, 81,* 135-146.

Lederman, E., & Blaine, J. R. 1972). Comparison of the level and predictive validity of Preschool Attainment Record ratings obtained from teachers and mothers. *Psychology in the Schools, 9,* 392-395.

Lewis, A. (1980). The early identification of children with learning difficulties. *Journal of Learning Disabilities, 13*(2), 51-57.

Lindsay, G. A., & Wedell, K. (1982). The early identification of educationally 'at risk' children revisited. *Journal of Learning Disabilities, 15*(4), 212-217.

Mulholland, J. J. (1972). Review of the Otis-Lennon Mental Ability Test. In O. K. Buros (ed.), *The seventh mental measurement yearbook.* Highland Park, N. J.: Gryphon Press.

Nurss, J. R. (1979). Assessment of readiness. In T. G. Waller & G. E. MacKinnon (eds.), *Reading research: Advances in theory and practice (vol. 1).* New York: Academic Press.

Otis, A. S., & Lennon, R. T. (1967). *Otis-Lennon Mental Ability Test*. New York: Harcourt, Brace & World.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction (2nd ed)*. New York: Holt, Rinehart & Winston.

Pope, J., Lehrer, B., & Stevens, J. (1980). A multiphasic reading screening procedure. *Journal of Reading Disabilities. 13* (2), 98-102.

Rubin, R. A., Barlow, B., Dorle, J., & Rosen, M. (1978). Preschool prediction of low achievement in basic school skills. *Journal of Learning Disabilities, 11*(10), 664-666.

Sattler, J. M. (1982). *Assessment of children's intelligence.* Philadelphia: W. B. Saunders.

Satz, P., & Friel, J. (1978). Predictive validity of an abbreviated screening battery. *Journal of Learning Disabilities, 11*(6), 347-350.

Service, J. (1972). *A user's guide to the Statistical Analysis System.* Raleigh, N. C.: North Carolina State University Press.

Smith. F. (1970). *Understanding reading.* New York: Holt, Rinehart & Winston.

Thompson, G. R. (1972). A comparison of social class effects in two tests of auditory discrimination. *Journal of Learning Disabilities,* 5(4), 37-41.